Nonlinear SDEs driven by Lévy processes and related PDEs

Benjamin Jourdain, Sylvie Méléard, Wojbor Woyczynski

Nonlinear SDEs driven by Lévy processes and related PDEs

Benjamin Jourdain*, Sylvie Méléard†, Wojbor A. Woyczynski‡

July 18, 2007

Abstract

In this paper we study general nonlinear stochastic differential equations, where the usual Brownian motion is replaced by a Lévy process. We also suppose that the coefficient multiplying the increments of this process is merely Lipschitz continuous and not necessarily linear in the time-marginals of the solution as is the case in the classical McKean-Vlasov model. We first study existence, uniqueness and particle approximations for these stochastic differential equations. When the driving process is a pure jump Lévy process with a smooth but unbounded Lévy measure, we develop a stochastic calculus of variations to prove that the time-marginals of the solutions are absolutely continuous with respect to the Lebesgue measure. In the case of a symmetric stable driving process, we deduce the existence of a function solution to a nonlinear integro-differential equation involving the fractional Laplacian.

Key words: Particle systems; Propagation of chaos; Nonlinear stochastic differential equations driven by Lévy processes; Partial differential equation with fractional Laplacian; Porous medium equation; McKean-Vlasov model.

MSC 2000: 60K35, 35S10, 65C35.

This paper studies the following nonlinear stochastic differential equation:

$$\begin{cases} X_t = X_0 + \int_0^t \sigma(X_{s^-}, P_s) dZ_s, \quad t \in [0,T], \\ \forall s \in [0,T], \ P_s \ \text{ denotes the probability distribution of } \ X_s. \end{cases} \tag{1}$$

We assume that X_0 is a random variable with values in $\mathbb{R}^k$, distributed according to m, $(Z_t)_{t\leq T}$ a Lévy process with values in $\mathbb{R}^d$, independent of X_0, and $\sigma : \mathbb{R}^k \times \mathcal{P}(\mathbb{R}^k) \to \mathbb{R}^{k\times d}$, where $\mathcal{P}(\mathbb{R}^k)$ denotes the set of probability measures on $\mathbb{R}^k$. Notice that the classical McKean-Vlasov model, studied for instance in [22], is obtained as a special case of (1) by choosing σ linear in the second variable and $Z_t = (t, B_t)$, with B_t being a $(d-1)$-dimensional standard Brownian motion.

The first section of the paper is devoted to the existence problem and particle approximations for (1). Initially, we address the case of square integrable both, the initial condition X_0, and the Lévy process $(Z_t)_{t\leq T}$. Under these assumptions the existence and uniqueness problem for (1) can be handled exactly as in the Brownian case $Z_t = (t, B_t)$. The nonlinear stochastic differential equation (1) admits a unique solution as soon as σ is Lipschitz continuous on $\mathbb{R}^k \times \mathcal{P}_2(\mathbb{R}^k)$ endowed with the product of the canonical metric on $\mathbb{R}^k$ and the Vaserstein metric d on the set $\mathcal{P}_2(\mathbb{R}^k)$ of probability measures with finite second order moments. This assumption is

*CERMICS, École des Ponts, ParisTech, 6-8 avenue Blaise Pascal, Cité Descartes, Champs sur Marne, 77455 Marne la Vallée Cedex 2, e-mail:jourdain@cermics.enpc.fr

†CMAP, Ecole Polytechnique, CNRS, route de Saclay, 91128 Palaiseau Cedex e-mail: sylvie.meleard@polytechnique.edu

‡Department of Statistics and Center for Stochastic and Chaotic Processes in Science and Technology, Case Western Reserve University, Cleveland, OH 44106, e-mail: waw@po.cwru.edu

much weaker than the assumptions imposed on σ in the classical McKean-Vlasov model, where it is also supposed to be linear in its second variable, that is, $\sigma(x,\nu) = \int_{\mathbb{R}^k} \varsigma(x,y)\nu(dy)$, for a Lipschitz continuous function $\varsigma : \mathbb{R}^k \times \mathbb{R}^k \to \mathbb{R}^{k\times d}$. Then, replacing the nonlinearity by the related interaction, we define systems of n interacting particles. In the limit $n \to +\infty$, we prove, by a trajectorial propagation of chaos result, that the dynamics of each particle approximates the one given by (1). Unlike in the very specific McKean-Vlasov model, where the universal $C/\sqrt{n}$ rate of convergence corresponds to the central limit theorem, under our general assumptions on σ, the rate of convergence turns out to depend on the spatial dimension k.

In the next step, the square integrability assumption is relaxed. However, to compensate for its loss, we assume a reinforced Lischitz continuity of σ : the Vaserstein metric d on $\mathcal{P}(\mathbb{R}^k)$ is replaced by its smaller and bounded modification d_1 defined below. Then, choosing square integrable approximants of the initial variable and the Lévy process, we prove existence for (1). Uniqueness remains an open question.

In the second section, we deal with the issue of absolute continuity of P_t when Z is a pure jump Lévy process with infinite intensity. For the sake of simplicity of the exposition, we restrict ourselves to the one-dimensional case $k = d = 1$. When σ does not vanish and admits two bounded derivatives with respect to its first variable, and the Lévy measure of Z satisfies some technical conditions, we prove that, for each $t > 0$, P_t has a density with respect to the Lebesgue measure on $\mathbb{R}$. The proof depends on a stochastic calculus of variations for the SDEs driven by Z which we develop by generalizing the approach of Bichteler-Jacod [4], (see also Bismut [5]), who dealt with the case of homogeneous processes with a jump measure equal to the Lebesgue measure. In our case, the nonlinearity induces an inhomogeneity in time and the jump measure is much more general, which introduces additional difficulties making the extension nontrivial. Graham-Méléard [11] developed similar techniques for a very specific stochastic differential equation related to the Kac equation. In that case, the jumps of the process were bounded. In our case, unbounded jumps are allowed and we deal with the resulting possible lack of integrability of the process X by an appropriate conditioning.

In the third section, we keep the assumptions made on σ in the second section, and assume that the driving Lévy process Z is symmetric and α-stable. Then, we apply the absolute continuity results obtained in Section 2 to prove that the solutions to (1) are such that for $t > 0$, P_t admits a density p_t with respect to the Lebesgue measure on the real line. In addition, calculating explicitly the adjoint of the generator of X, we conclude that the function $p_t(x)$ is a weak solution to the nonlinear Fokker-Planck equation

$$\begin{cases} \partial_t p_t(x) = D_x^\alpha(|\sigma(.,p_t)|^\alpha p_t(.))(x) \\ \lim_{t\to 0^+} p_t(x)dx = m(dx), \end{cases},$$

where, by a slight abuse of the notation, $\sigma(.,p_t)$ stands for $\sigma(.,p_t(y)dy)$, the limit is understood in the sense of the narrow convergence, and $D_x^\alpha = -(-\Delta)^{\alpha/2}$ denotes the spatial, spherically symmetric fractional derivative of order α defined here as a singular integral operator,

$$D_x^\alpha f(x) = K \int_{\mathbb{R}} \left(f(x+y) - f(x) - \mathbf{1}_{\{|y|\le 1\}} f'(x)y\right) \frac{dy}{|y|^{1+\alpha}},$$

where K is a positive constant. For

$$\sigma(x,\nu) = (g_\varepsilon * \nu(x))^s \text{ with } \varepsilon > 0,\ g_\varepsilon(x) = \frac{1}{\sqrt{2\pi\varepsilon}} e^{-\frac{x^2}{2\varepsilon}} \text{ and } s > 0,$$

one obtains the nonlocal approximation $\partial_t p_t = D_x^\alpha((g_\varepsilon * p_t)^{\alpha s} p_t)$ of the fractional porous medium equation $\partial_t p_t = D_x^\alpha(p_t^{\alpha s+1})$, the physical interest of which is discussed at the end of the paper.

Other nonlinear evolution equations involving generators of Lévy processes, such as fractional conservation laws have been studied via probabilistic tools in, e.g., [13], and[14].

Notations : Throughout the paper, C will denote a constant which may change from line to line. In spaces with finite dimension, the Euclidian norm is denoted by $|\ \ |$. Let $\mathcal{P}(\mathbb{R}^k)$ denote the set of probability measures on $\mathbb{R}^k$, and $\mathcal{P}_2(\mathbb{R}^k)$ – the subset of measures with finite second order moments. For $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^k)$, the *Vaserstein metric* is defined by the formula,

$$d(\mu,\nu) = \inf\left\{ \left(\int_{\mathbb{R}^k\times\mathbb{R}^k} |x-y|^2\, Q(dx,dy) \right)^{1/2} : Q \in \mathcal{P}(\mathbb{R}^k\times\mathbb{R}^k) \text{ with marginals } \mu \text{ and } \nu \right\}.$$

It induces the topology of weak convergence together with convergence of moments up to order 2. The modified Vaserstein metric on $\mathcal{P}(\mathbb{R}^k)$ defined by the formula,

$$d_1(\mu,\nu) = \inf\left\{ \left(\int_{\mathbb{R}^k\times\mathbb{R}^k} |x-y|^2 \wedge 1\, Q(dx,dy) \right)^{1/2} : Q \in \mathcal{P}(\mathbb{R}^k\times\mathbb{R}^k) \text{ with marginals } \mu \text{ and } \nu \right\},$$

simply induces the topology of weak convergence.

1 Existence of a nonlinear process

We first address the case when both, the initial condition X_0 and Z are square integrable, before relaxing these integrability conditions later on.

1.1 The square integrable case

In this subsection we assume that the initial condition X_0, and the Lévy process $(Z_t)_{t\le T}$, are both square integrable : $\mathbb{E}(|X_0|^2+|Z_T|^2) < +\infty$. Under this assumption, the following inequality generalizes the Brownian case (see, [20], Theorem 66, p.339) :

Lemma 1 *Let $p \ge 2$ be such that $\mathbb{E}(|Z_T|^p) < +\infty$. There is a constant C_p such that, for any $\mathbb{R}^{k\times d}$-valued process $(H_t)_{t\le T}$ predictable for the filtration $\mathcal{F}_t = \sigma(X_0, (Z_s)_{s\le t})$, $\forall t \in [0,T]$,*

$$\mathbb{E}\left(\sup_{s\le t} \left| \int_0^s H_u dZ_u \right|^p \right) \le C_p \int_0^t \mathbb{E}(|H_s|^p) ds.$$

Because of this inequality for $p = 2$, the results obtained for the classical McKean-Vlasov model driven by a standard Brownian motion still hold. First, we state and prove

Proposition 2 *Assume that X_0 and $(Z_t)_{t\le T}$ are square integrable, and that the mapping σ is Lipschitz continuous when $\mathbb{R}^k \times \mathcal{P}_2(\mathbb{R}^k)$ is endowed with the product of the canonical topology on $\mathbb{R}^k$ and the Vaserstein metric d on $\mathcal{P}_2(\mathbb{R}^k)$. Then equation* (1) *admits a unique solution such that $\mathbb{E}\left(\sup_{t\le T} |X_t|^2\right) < +\infty$. Moreover, if for some $p > 2$, $\mathbb{E}(|X_0|^p + |Z_T|^p) < +\infty$, then $\mathbb{E}\left(\sup_{t\le T} |X_t|^p\right) < +\infty$.*

Proof : We generalize here the pathwise fixed point approach well known in the classical McKean-Vlasov case (see, Sznitman [22]). Let $\mathbb{D}$ denote the space of càdlàg functions from $[0,T]$

to $\mathbb{R}^k$, and $\mathcal{P}_2(\mathbb{D})$ the space of probability measures Q on $\mathbb{D}$ such that $\int_{\mathbb{D}} \sup_{t\le T} |Y_t|^2 Q(dY) < +\infty$. Endowed with the Vaserstein metric $D_T(P,Q)$ where, for $t \le T$,

$$D_t(P,Q) = \inf\left\{ \left(\int_{\mathbb{D}\times\mathbb{D}} \sup_{s\le t} |Y_s - W_s|^2 \, R(dY, dW) \right)^{1/2} : R \in \mathcal{P}(\mathbb{D}\times\mathbb{D}) \text{ with marginals } P \text{ and } Q \right\},$$

$\mathcal{P}_2(\mathbb{D})$ is a complete space.

For $Q \in \mathcal{P}(\mathbb{D})$ with time-marginals $(Q_t)_{t\in[0,T]}$, in view of Lebesgue's Theorem, the distance

$$d(Q_t, Q_s) \le \int_{\mathbb{D}} |Y_t - Y_s|^2 Q(dY)$$

converges to 0, as s decreases to t (respectively, $d(Q_{t^-}, Q_s) \le \int_{\mathbb{D}} |Y_{t^-} - Y_s|^2 Q(dY)$ converges to 0, as s increases to t; here $Q_{t^-} = Q \circ Y_{t^-}^{-1}$ is the weak limit of Q_s as $s \to t^-$). Therefore, the mapping $t \in [0,T] \to Q_t$ is càdlàg when $\mathcal{P}_2(\mathbb{R}^k)$ is endowed with the metric d. As a consequence, for fixed $x \in \mathbb{R}^k$, the mapping $t \in [0,T] \to \sigma(x, Q_t)$ is càdlàg. Hence, by a multidimensional version of Theorem 6, p. 249, in [20], the standard stochastic differential equation

$$X_t^Q = X_0 + \int_0^t \sigma(X_{s^-}^Q, Q_s) dZ_s, \; t \in [0,T]$$

admits a unique solution.

Let Φ denote the mapping on $\mathcal{P}_2(\mathbb{D})$ which associates the law of X^Q with Q. Let us check that Φ takes its values in $\mathcal{P}_2(\mathbb{D})$. For $K > 0$, we set $\tau_K = \inf\{s \le T : |X_s^Q| \ge K\}$. By Lemma 1 and the Lipschitz property of σ, one has

$$\begin{aligned}
\mathbb{E}\left(\sup_{s\le t} |X_{s\wedge\tau_K}^Q|^2 \right) &\le C \left(\mathbb{E}(|X_0|^2) + \int_0^t \mathbb{E}\left(1_{\{s\le\tau_K\}} |\sigma(X_s^Q, Q_s) - \sigma(0,\delta_0)|^2 + |\sigma(0,\delta_0)|^2 \right) ds \right) \\
&\le C \left(\mathbb{E}(|X_0|^2) + \int_0^t \mathbb{E}\left(\sup_{r\le s} |X_{r\wedge\tau_K}^Q|^2 \right) ds + t \int_{\mathbb{D}} \sup_{t\le T} |Y_t|^2 Q(dY) + t|\sigma(0,\delta_0)|^2 \right).
\end{aligned}$$

By Gronwall's Lemma, one deduces that

$$\mathbb{E}\left(\sup_{s\le T} |X_{s\wedge\tau_K}^Q|^2 \right) \le C \left(\mathbb{E}(|X_0|^2) + |\sigma(0,\delta_0)|^2 + \int_{\mathbb{D}} \sup_{t\le T} |Y_t|^2 Q(dY) \right),$$

where the constant C does not depend on K. Letting K tend to $+\infty$, one concludes by Fatou's Lemma that

$$\int_{\mathbb{D}} \sup_{s\le T} |Y_s|^2 d\Phi(Q)(Y) = \mathbb{E}\left(\sup_{s\le T} |X_s^Q|^2 \right) \le C \left(\mathbb{E}(|X_0|^2) + |\sigma(0,\delta_0)|^2 + \int_{\mathbb{D}} \sup_{t\le T} |Y_t|^2 Q(dY) \right). \tag{2}$$

Observe that a process $(X_t)_{t\in[0,T]}$, such that $\mathbb{E}\left(\sup_{t\le T} |X_t|^2\right) < +\infty$, solves (1) if and only if its law is a fixed-point of Φ. So, to complete the proof of the Proposition, it suffices to check that Φ admits a unique fixed point.

By a formal computation, which can be made rigorous by a localization procedure similar to the one utilized above, for $P, Q \in \mathcal{P}_2(\mathbb{D})$ one has

$$\begin{aligned}
\mathbb{E}\left(\sup_{s\le t} |X_s^P - X_s^Q|^2 \right) &\le C \int_0^t \mathbb{E}(|\sigma(X_{s^-}^P, P_s) - \sigma(X_{s^-}^Q, Q_s)|^2) ds \\
&\le C \int_0^t \mathbb{E}\left(\sup_{r\le s} |X_r^P - X_r^Q|^2 \right) + d^2(P_s, Q_s) ds.
\end{aligned}$$

By Gronwall's Lemma, one deduces that, $\forall t \leq T$,

$$\mathbb{E}\left(\sup_{s\leq t}|X_s^P - X_s^Q|^2\right) \leq C\int_0^t d^2(P_s,Q_s)ds.$$

Since $D_t^2(\Phi(P),\Phi(Q)) \leq \mathbb{E}(\sup_{s\leq t}|X_s^P - X_s^Q|^2)$, and $d(P_s,Q_s) \leq D_s(P,Q)$, the last inequality implies, $\forall t \leq T$,

$$D_t^2(\Phi(P),\Phi(Q)) \leq C\int_0^t D_s^2(P,Q)ds.$$

Iterating this inequality, and denoting by Φ^N the N-fold composition of Φ, we obtain that, $\forall N \in \mathbb{N}^*$,

$$D_T^2(\Phi^N(P),\Phi^N(Q)) \leq C^N\int_0^T \frac{(T-s)^{N-1}}{(N-1)!}D_s^2(P,Q)ds \leq \frac{C^N T^N}{N!}D_T^2(P,Q).$$

Hence, for N large enough, Φ^N is a contraction which entails that Φ admits a unique fixed point.

If, for some $p > 2$, $\mathbb{E}(|X_0|^p + |Z_T|^p) < +\infty$, a reasoning similar to the one used in the derivation of (2), easily leads to the conclusion that the constructed solution $(X_t)_{t\leq T}$ of (1) is such that

$$\mathbb{E}\left(\sup_{s\leq T}|X_s|^p\right) \leq C\left(\mathbb{E}(|X_0|^p) + |\sigma(0,\delta_0)|^p + \mathbb{E}\left(\sup_{s\leq T}|X_s|^2\right)^{p/2}\right) < +\infty.$$

■

Our next step is to study pathwise particle approximations for the nonlinear process. Let $((X_0^i, Z^i))_{i\in\mathbb{N}^*}$ denote a sequence of independent pairs with (X_0^i, Z^i) distributed like (X_0, Z). For each $i \geq 1$, let $(X_t^i)_{t\in[0,T]}$ denote the solution given by Proposition 2 of the nonlinear stochastic differential equation starting from X_0^i and driven by Z^i :

$$\begin{cases} X_t^i = X_0^i + \int_0^t \sigma(X_{s^-}^i, P_s)dZ_s^i,\ t\in[0,T] \\ \forall s\in[0,T],\ P_s \text{ denotes the probability distribution of } X_s^i \end{cases}. \tag{3}$$

Replacing the nonlinearity by interaction, we introduce the following system of n interacting particles

$$\begin{cases} X_t^{i,n} = X_0^i + \int_0^t \sigma(X_{s^-}^{i,n}, \mu_{s^-}^n)dZ_s^i, \quad t\in[0,T], \quad 1\leq i\leq n, \\ \text{where } \mu^n = \frac{1}{n}\sum_{j=1}^n \delta_{X^{j,n}} \text{ denotes the empirical measure} \end{cases}. \tag{4}$$

Since for $\xi = (x_1,\ldots,x_n)$, and $\zeta = (y_1,\ldots,y_n)$ in $\mathbb{R}^{nk}$, one has

$$d\left(\frac{1}{n}\sum_{j=1}^n \delta_{x_j}, \frac{1}{n}\sum_{j=1}^n \delta_{y_j}\right) \leq \left(\frac{1}{n}\sum_{j=1}^n |x_j - y_j|^2\right)^{1/2} = \frac{1}{\sqrt{n}}|\xi - \zeta|. \tag{5}$$

Existence of a unique solution to (4), with finite second order moments, follows from Theorem 7, p. 253, in [20]. Our next result establishes the trajectorial propagation of chaos result for the interacting particle system (4).

Theorem 3 *Under the assumptions of Proposition 2,*

$$\lim_{n\to+\infty} \sup_{i\le n} \mathbb{E}\left(\sup_{t\le T} |X_t^{i,n} - X_t^i|^2\right) = 0$$

Moreover, under additional assumptions, the following two explicit estimates hold :

- *If* $\mathbb{E}(|X_0|^{k+5} + |Z_T|^{k+5}) < +\infty$, *then*

$$\sup_{i\le n} \mathbb{E}\left(\sup_{t\le T} |X_t^{i,n} - X_t^i|^2\right) \le C n^{-\frac{2}{k+4}}; \tag{6}$$

- *If* $\sigma(x,\nu) = \int_{\mathbb{R}^k} \varsigma(x,y)\nu(dy)$, *where* $\varsigma : \mathbb{R}^k \times \mathbb{R}^k \to \mathbb{R}^{k\times d}$ *is a Lipschitz continuous function, then*

$$\sup_{i\le n} \mathbb{E}\left(\sup_{t\le T} |X_t^{i,n} - X_t^i|^2\right) \le \frac{C}{n}, \tag{7}$$

  where the constant C does not depend on n.

The proof of the first assertion relies on the following

Lemma 4 *Let ν be a probability measure on $\mathbb{R}^k$ such that $\int_{\mathbb{R}^k} |x|^2\nu(dx) < +\infty$, and $\nu^n = \frac{1}{n}\sum_{j=1}^n \delta_{\xi_j}$ denote the empirical measure associated with a sequence $(\xi_i)_{i\ge1}$ of independent random variables with law ν. Then, $\forall n \ge 1$,*

$$\mathbb{E}\left(d^2\left(\nu^n,\nu\right)\right) \le 4\int_{\mathbb{R}^k} |x|^2\nu(dx), \quad \textit{and} \quad \lim_{n\to+\infty} \mathbb{E}\left(d^2\left(\nu^n,\nu\right)\right) = 0.$$

Proof of Lemma 4 : By the strong law of large numbers, as n tends to ∞, almost surely ν^n converges weakly to ν and, $\forall i,j \in \{1,\ldots,k\}$, $\int_{\mathbb{R}^k} x_i\nu_n(dx)$ (resp. $\int_{\mathbb{R}^k} x_ix_j\nu_n(dx)$) converges to $\int_{\mathbb{R}^k} x_i\nu(dx)$ (resp. $\int_{\mathbb{R}^k} x_ix_j\nu(dx)$). Since the Vaserstein distance d induces the topology of simultaneous weak convergence and convergence of moments up to order 2, one deduces that almost surely, $d(\nu^n,\nu)$ converges to 0, as n tends to ∞. Hence, to conclude the proof of the first assertion, it is enough to check that the random variables $(d^2(\nu^n,\nu))_{n\ge1}$ are uniformly integrable. To see that note the inequality

$$d^2(\nu^n,\nu) \le \frac{2}{n}\sum_{j=1}^n |\xi_j|^2 + 2\int_{\mathbb{R}^k} |x|^2\nu(dx).$$

The right-hand side is nonnegative and converges almost surely to $4\int_{\mathbb{R}^k} |x|^2\nu(dx)$, as $n\to\infty$. Since its expectation is constant, and equal to the expectation of the limit, one deduces that the convergence is also in L^1. As a consequence, for $n \ge 1$, the random variables in the right-hand side, and therefore in the left-hand side, are uniformly integrable. ∎

Proof of Theorem 3 : Let $P^n = \frac{1}{n}\sum_{j=1}^n \delta_{X^j}$ denote the empirical measure of the independent nonlinear processes (3). By a formal computation, which can be made rigorous by a localization argument similar to the one made in the proof of Proposition 2, one has, $\forall t \leq T$,

$$\begin{aligned}\mathbb{E}\left(\sup_{s\leq t}|X_s^{i,n} - X_s^i|^2\right) \leq & C\int_0^t \mathbb{E}\left(|\sigma(X_s^{i,n},\mu_s^n) - \sigma(X_s^i,P_s^n)|^2\right) ds \\ & + C\int_0^t \mathbb{E}\left(|\sigma(X_s^i,P_s^n) - \sigma(X_s^i,P_s)|^2\right) ds.\end{aligned}$$

In view of the Lipschitz property of σ, the estimate (5), and the exchangeability of the couples $(X^i, X^{i,n})_{1\leq i\leq n}$, the first term of the right is smaller than $C\int_0^t \mathbb{E}\left(\sup_{r\leq s}|X_r^{i,n} - X_r^i|^2\right) ds$. By Gronwall's Lemma, and the Lipschitz assumption on σ, one deduces that

$$\mathbb{E}\left(\sup_{t\leq T}|X_t^{i,n} - X_t^i|^2\right) \leq C\int_0^T \mathbb{E}\left(\left|\sigma\left(X_s^i,P_s^n\right) - \sigma(X_s^i,P_s)\right|^2\right) ds \leq C\int_0^T \mathbb{E}(d^2(P_s^n,P_s))ds.$$

The first assertion then follows from Lemma 4, the upper-bounds of the second order moments given in Proposition 2, and by Lebesgue's Theorem.

The second assertion is deduced from the upper-bounds for moments of order $k+5$ combined with the following restatement of Theorem 10.2.6 in [21] :

$$\mathbb{E}\left(d^2\left(P_s^n,P_s\right)\right) \leq C\left(1 + \sqrt{\int_{\mathbb{R}^k}|y|^{k+5}P_s(dy)}\right) n^{-\frac{2}{k+4}},$$

where the constant C only depends on k. The precise dependence of the upper-bound on $\int_{\mathbb{R}^k}|y|^{k+5}P_s(dy)$ comes from a carefull reading of the proof given in [21].

Finally, if, as in the usual McKean-Vlasov framework (see [22]), $\sigma(x,\nu) = \int_{\mathbb{R}^k}\varsigma(x,y)\nu(dy)$, where $\varsigma = (\varsigma_{ab})_{a\leq k, b\leq d} : \mathbb{R}^k\times\mathbb{R}^k \to \mathbb{R}^{k\times d}$ is Lipschitz continuous, then $\mathbb{E}\left(\left|\sigma\left(X_s^i,P_s^n\right) - \sigma(X_s^i,P_s)\right|^2\right)$ is equal to

$$\sum_{a=1}^k\sum_{b=1}^d\frac{1}{n^2}\sum_{j,l=1}^n \mathbb{E}\left(\left[\varsigma_{ab}(X_s^i,X_s^j) - \int_{\mathbb{R}^k}\varsigma_{ab}(X_s^i,y)P_s(dy)\right]\left[\varsigma_{ab}(X_s^i,X_s^l) - \int_{\mathbb{R}^k}\varsigma_{ab}(X_s^i,y)P_s(dy)\right]\right).$$

Since, by independence of the random variables $X_s^1,\ldots,X_s^n$ with common law P_s, the expectation in the above summation vanishes as soon as $j \neq l$, the third assertion of Theorem 3 easily follows. ∎

Remark 5

- Observe the lower estimate

$$d(\nu^n,\nu) \geq \left(\int_{\mathbb{R}^k}\min_{1\leq j\leq n}|\xi_j - x|^2\nu(dx)\right)^{1/2} \geq \inf_{(y_1,\ldots,y_n)\in(\mathbb{R}^k)^n}\left(\int_{\mathbb{R}^k}\min_{1\leq j\leq n}|y_j - x|^2\nu(dx)\right)^{1/2}.$$

  Moreover, according to the Bucklew and Wise Theorem [6], if ν has a density φ with respect to the Lebesgue measure on $\mathbb{R}^k$ which belongs to $L^{\frac{k}{2+k}}(\mathbb{R}^k)$, then, as n tends to infinity,

$$n^{1/k}\inf_{(y_1,\ldots,y_n)\in(\mathbb{R}^k)^n}\left(\int_{\mathbb{R}^k}\min_{1\leq j\leq n}|y_j - x|^2\varphi(x)dx\right)^{1/2}$$

converges to $C_k\|\varphi\|_{\frac{k}{2+k}}$, where the constant C_k only depends on k. Hence, one cannot expect $\mathbb{E}(d^2(\nu^n,\nu))$ to vanish quicker than $Cn^{-2/k}$.

Therefore, if $\nu \to \sigma(x,\nu)$ is merely Lipschitz continuous for the Vaserstein metric, one cannot expect $\mathbb{E}\left(\sup_{t\leq T}|X_t^{i,n} - X_t^i|^2\right)$ to vanish quicker than $Cn^{-2/k}$. The rate of convergence obtained in (6) is not far from being optimal at least for a large spatial dimension k. Nevertheless, in the McKean-Vlasov framework, where the structure of σ is very specific, one can overcome this dependence of the convergence rate on the dimension k, and recover the usual central limit theorem rate.

- The square integrability assumption on the initial variable X_0 can be relaxed if σ is Lipschitz continuous with $\mathbb{R}^k \times \mathcal{P}(\mathbb{R}^k)$ endowed with the product of the canonical topology on $\mathbb{R}^k$ and the modified Vaserstein metric d_1 on $\mathcal{P}(\mathbb{R}^k)$. Indeed, one may then adapt the fixed-point approach in the proof of Proposition 2 by defining $\mathcal{P}$ as the space of probability measures on $\mathbb{D}$, and replacing $\sup_{s\leq t}|Y_s - W_s|^2$ by $\sup_{s\leq t}|Y_s - W_s|^2 \wedge 1$ in the definition of $D_t(P,Q)$. This way, one obtains that the nonlinear stochastic differential equation (1) still admits a unique solution even if the initial condition X_0 is not square integrable. Moreover, for any probability measure ν on $\mathbb{R}^k$, if ν_n is defined as above, $\mathbb{E}(d_1^2(\nu_n,\nu))$ remains bounded by one, and converges to 0 as n tends to infinity. Therefore, the first assertion in Theorem 3 still holds.

The next subsection is devoted to the more complicated case when the square integrability assumption on the Lévy process $(Z_t)_{t\leq T}$ is also relaxed.

1.2 The general case

In this section, we impose no integrability conditions, either on the initial condition X_0, or on the Lévy process $(Z_t)_{t\leq T}$. Let us denote by $m \in \mathcal{P}(\mathbb{R}^k)$ the distribution of the former. According to the Lévy-Khintchine formula, the infinitesimal generator of the latter can be written, for $f \in C_b^2(\mathbb{R}^d)$, in the form

$$Lf(z) = \frac{1}{2}\sum_{i,j=1}^d a_{ij}\partial^2_{z_i,z_j}f(z) + b.\nabla f(z) + \int_{\mathbb{R}^d}\left[f(z+y) - f(z) - \mathbf{1}_{\{|y|\leq 1\}}y.\nabla f(z)\right]\beta(dy),$$

where $a = (a_{ij})_{1\leq i,j\leq d}$ is a non-negative symmetric matrix, b a given vector in $\mathbb{R}^d$, and β a measure on $\mathbb{R}^d$ satisfying the integrability condition $\int_{\mathbb{R}^d}(1\wedge|y|^2)\beta(dy) < +\infty$.

To deal with non square integrable sources of randomness, we impose a stronger continuity condition on σ, namely, we assume that σ is Lipschitz continuous when $\mathbb{R}^k \times \mathcal{P}(\mathbb{R}^k)$ is endowed with the product of the canonical topology on $\mathbb{R}^k$ and the modified Vaserstein metric d_1 on $\mathcal{P}(\mathbb{R}^k)$. Notice that, under this assumption, for each $x \in \mathbb{R}^k$, the mapping $\nu \in \mathcal{P}(\mathbb{R}^k) \to \sigma(x,\nu)$ is bounded.

In order to prove existence of a weak solution to (1), we introduce a cutoff parameter $N \in \mathbb{N}^*$, and define a square integrable initial random variable $X_0^N = X_0 1_{\{|X_0|\leq N\}}$, and a square integrable Lévy process $(Z_t^N)_{t\leq T}$ by removing the jumps of $(Z_t)_{t\leq T}$ larger than N :

$$Z_t^N = Z_t - \sum_{s\leq t} 1_{\{|\Delta Z_s|>N\}}\Delta Z_s.$$

Let $(X_t^N)_{t\in[0,T]}$ denote the solution given by Proposition 2 of the nonlinear stochastic differential equation starting from X_0^N and driven by $(Z_t^N)_{t\in[0,T]}$:

$$\begin{cases} X_t^N = X_0^N + \int_0^t \sigma(X_{s^-}^N, P_s^N) dZ_s^N, \ t \in [0,T] \\ \forall s \in [0,T], \ P_s^N \ \text{ denotes the probability distribution of } \ X_s^N \end{cases}. \tag{8}$$

We are going to prove that when the cutoff parameter N tends to ∞, then $(X_t^N)_{t\in[0,T]}$ converges in law to a weak solution of (1). More precisely, let us denote by P^N the distribution of $(X_t^N)_{t\in[0,T]}$, and by $(Y_t)_{t\in[0,T]}$ the canonical process on $\mathbb{D}$.

Proposition 6 *The set of probability measures $(P^N)_{N\in\mathbb{N}^*}$ is tight when $\mathbb{D}$ is endowed with the Skorohod topology. In addition, any weak limit P, with time marginals $(P_t)_{t\in[0,T]}$, of its converging subsequences solves the following martingale problem :*

$$\begin{cases} P_0 = m \ \text{ and } \forall \varphi : \mathbb{R}^k \to \mathbb{R}, \ C^2 \text{ with compact support}, \\ \left(M_t^\varphi = \varphi(Y_t) - \varphi(Y_0) - \int_0^t \mathcal{L}[P_s]\varphi(Y_s)ds \right)_{t\in[0,T]} \ \text{ is a } P\text{-martingale} \end{cases}, \tag{9}$$

where for each $\nu \in \mathcal{P}(\mathbb{R}^k)$, and any $x \in \mathbb{R}^k$,

$$\begin{aligned} \mathcal{L}[\nu]\varphi(x) =& \frac{1}{2}\sum_{i,j=1}^k (\sigma a \sigma^*(x,\nu))_{ij} \partial^2_{x_i,x_j}\varphi(x) + (\sigma(x,\nu)b).\nabla\varphi(x) \\ &+ \int_{\mathbb{R}^d} \left[\varphi\left(x + \sigma(x,\nu)y\right) - \varphi(x) - \mathbf{1}_{\{|y|\le 1\}}\sigma(x,\nu)y.\nabla\varphi(x)\right] \beta(dy)ds. \end{aligned} \tag{10}$$

Proof : Let us first remark that for $N \in \mathbb{N}^*$, and for a fixed $x \in \mathbb{R}^k$, the mapping $t \in [0,T] \to \sigma(x, P_t^N)$ is càdlàg and bounded by a constant not depending on N. As a consequence, according to Theorem 6 p. 249 [20], for a fixed $M \in \mathbb{N}^*$, the stochastic differential equation

$$X_t^{N,M} = X_0^{N\wedge M} + \int_0^t \sigma(X_{s-}^{N,M}, P_s^N) dZ_s^{N\wedge M}, \quad t \in [0,T],$$

admits a unique solution. Let us denote by $P^{N,M}$ the law of $(X_t^{N,M})_{t\in[0,T]}$. By trajectorial uniqueness, $\forall N \in \mathbb{N}^*$, $\forall t \in [0,T]$,

$$X_t^N = X_t^{N,M},$$

as long as $|X_0| \vee \sup_{t\in]0,T]} |\Delta Z_t| \le M$. The probability of the latter event tends to one as M tends to infinity. Using both the necessary and the sufficient conditions of Prokhorov's Theorem, one deduces that the tightness of the sequence $(P^N)_{N\in\mathbb{N}^*}$ is implied by the tightness of the sequence $(P^{N,M})_{N\in\mathbb{N}^*}$, for any fixed $M \in \mathbb{N}^*$.

Let us now prove this last result by fixing $M \in \mathbb{N}^*$. Using the boundedness of d_1, one easily checks that

$$\sup_{N\in\mathbb{N}^*} \mathbb{E}\left(\sup_{t\le T} |X_t^{N,M}|^2 \right) < +\infty. \tag{11}$$

This implies tightness of the laws of the random variables $\left(\sup_{t\le T} |X_t^{N,M}|\right)_{N\in\mathbb{N}^*}$. In order to use Aldous' criterion, we set $\varepsilon, \delta > 0$, and introduce two stopping times S, and $\tilde{S}$, such that $0 \le S \le \tilde{S} \le (S+\delta)\wedge T$. Let us also remark that, for $K \in \mathbb{N}^*$, and $b^K = b + \int_{\mathbb{R}^d} y \mathbf{1}_{\{1<|y|\le K\}}\beta(dy)$, the process $\left(\tilde{Z}_t^K = Z_t^K - b^K t\right)_{t\in[0,T]}$ is a centered Lévy process and therefore a martingale. Now, observe that

$$\mathbb{P}\left(|X_{\tilde{S}}^{N,M} - X_S^{N,M}|^2 \geq \varepsilon\right) \leq \mathbb{P}\left(\left|\int_S^{\tilde{S}} \sigma(X_s^{N,M}, P_s^N) b^{N\wedge M} ds\right|^2 \geq \frac{\varepsilon}{4}\right) + \mathbb{P}\left(\left|\int_S^{\tilde{S}} \sigma(X_{s^-}^{N,M}, P_s^N) d\tilde{Z}_s^{N\wedge M}\right|^2 \geq \frac{\varepsilon}{4}\right). \quad (12)$$

Using the boundedness of the sequence $(b^{N\wedge M})_{N\in\mathbb{N}^*}$, the Lipschitz property of σ with respect to its first variable, and (11) combined with the inequalities of Markov and Cauchy-Schwarz, one obtains that the first term of the right-hand-side is smaller than $C\delta^2/\varepsilon$, where the constant C does not depend on N. For the second term of the right-hand-side, one remarks that Doob's optional sampling Theorem, followed by the Lipschitz property of σ, and (11), imply that

$$\begin{aligned}\mathbb{E}\left(\left|\int_S^{\tilde{S}} \sigma(X_{s^-}^{N,M}, P_s^N) d\tilde{Z}_s^{N\wedge M}\right|^2\right) &= \mathbb{E}\Bigg(\int_S^{\tilde{S}} \Bigg[\sum_{i=1}^k (\sigma a \sigma^*)_{ii}(X_s^{N,M}, P_s^N) \\ &\quad + \int_{\mathbb{R}^d} |\sigma(X_s^{N,M}, P_s^N) y|^2 \mathbf{1}_{\{|y|\leq N\wedge M\}} \beta(dy)\Bigg] ds\Bigg) \\ &\leq C\delta,\end{aligned}$$

where C does not depend on N. By Markov's Inequality, the second term of the right-hand-side of (12) is smaller than $C\delta/\varepsilon$ and, in view of Aldous' criterion, we conclude that the sequence $(P^{N,M})_{N\in\mathbb{N}^*}$ is tight.

Now, let us denote by P the limit of a converging subsequence of $(P^N)_{N\in\mathbb{N}^*}$ that we still index by N for simplicity's sake. Also, let φ denote a compactly supported C^2 function on $\mathbb{R}^k$. For $p \in \mathbb{N}^*$, $0 \leq s_1 \leq s_2 \leq \ldots \leq s_p \leq s \leq t \leq T$, and a continuous and bounded function $\psi : (\mathbb{R}^k)^p \to \mathbb{R}$, let F denote the mapping on $\mathcal{P}(\mathbb{D})$ defined by

$$F(Q) = \int_{\mathbb{D}} \left(\varphi(Y_t) - \varphi(Y_s) - \int_s^t \mathcal{L}[Q_u]\varphi(Y_u) du\right) \psi(Y_{s_1}, \ldots, Y_{s_p}) Q(dY).$$

For F^N defined like F, but with $\mathbf{1}_{\{|y|\leq N\}}\beta(dy)$ replacing $\beta(dy)$ in the definition (10) of $\mathcal{L}[\nu]$, one has $F^N(P^N) = 0$. Therefore

$$|F(P^N)| = |F(P^N) - F^N(P^N)| \leq 2(t-s)\|\psi\|_\infty \|\varphi\|_\infty \int_{\mathbb{R}^d} \mathbf{1}_{\{|y|\geq N\}} \beta(dy) \overset{N\to+\infty}{\longrightarrow} 0.$$

The mapping $(x, \nu) \in \mathbb{R}^k \times \mathcal{P}(\mathbb{R}^k) \to \mathcal{L}[\nu]\varphi(x)$ is bounded, continuous in x for a fixed ν, and continuous in ν, uniformly for $x \in \mathbb{R}^k$. Therefore, as soon as $s_1, \ldots, s_p, s, t$ do not belong to the at most countable set $\{u \in]0, T] :\ P(\Delta Y_u \neq 0) > 0\}$, then F is continuous and bounded at point P which implies $F(P) = \lim_{N\to+\infty} F(P^N) = 0$. In view of the right continuity of $u \to Y_u$ and Lebesgue's theorem, this equality still holds without any restriction on $s_1, \ldots, s_p, s, t$. Hence, $(M_t^\varphi)_{t\in[0,T]}$ is a P-martingale. Since the sequence $(X_0^N)_{N\in\mathbb{N}^*}$ converges in distribution to X_0, $P_0 = m$, which concludes the proof. ∎

The above existence result for the martingale problem (9) implies an analogous existence statement for the corresponding nonlinear Fokker-Planck equation.

Proposition 7 *Let P denote a solution of* (9). *Then the time marginals $(P_t)_{t\in[0,T]}$ solve the initial value problem*

$$\partial_t P_t = \mathcal{L}^*[P_t]P_t, \quad P_0 = m, \tag{13}$$

in the weak sense, where, for $\nu \in \mathcal{P}(\mathbb{R}^k)$, $\mathcal{L}^*[\nu]$ denotes the formal adjoint of $\mathcal{L}[\nu]$ defined by the following condition: $\forall \phi, \psi$ C^2 with compact support on $\mathbb{R}^k$,

$$\int_{\mathbb{R}^k} \mathcal{L}^*[\nu]\psi(x)\varphi(x)dx = \int_{\mathbb{R}^k} \psi(x)\mathcal{L}[\nu]\varphi(x)dx.$$

Moreover, the standard stochastic differential equation

$$X_t^P = X_0 + \int_0^t \sigma(X_{s^-}^P, P_s)dZ_s \tag{14}$$

admits a unique solution $(X_t^P)_{t\in[0,T]}$ and, for each $t \in [0,T]$, X_t^P is distributed according to P_t.

Proof : The first assertion follows readily from the constancy of the expectation of the P-martingale $(M_t^\varphi)_{t\in[0,T]}$. Existence and uniqueness for (14) follows from [20], Theorem 6 p. 249. Now, if Q_t denotes the law of X_t^P for $t \geq 0$, then $(Q_t)_{t\geq 0}$ solves

$$\partial_t Q_t = \mathcal{L}^*[P_t]Q_t, \quad Q_0 = m,$$

in the weak sense. Since $(P_t)_{t\geq 0}$ also is a weak solution of this linear equation, by Theorem 5.2 [3], one concludes that $\forall t \leq T$, $P_t = Q_t$. ∎

Remark 8 We have not been able to prove uniqueness for the nonlinear martingale problem (9). However, our assumptions, and Theorem 6, p.249, in [20], ensure existence and uniqueness for the particle system (4). Like in the proof of Proposition 6, one can check that the laws of the processes $X^{1,n}$, $n \geq 1$, are tight. According to [22], this implies uniform tightness of the laws π_n of the empirical measures μ^n. For a fixed $x \in \mathbb{R}^k$, the function $\nu \to \sigma(x,\nu)$ is continuous and bounded when $\mathcal{P}(\mathbb{R}^k)$ is endowed with the weak convergence topology. Then one can prove that the limit points of the sequence $(\pi_n)_n$ give full weight to the solutions of the nonlinear martingale problem (9).

2 Absolute continuity of the marginals

In this section we restrict ourselves to the one-dimensional case $k = d = 1$, and assume that Z is a pure jump Lévy process with a Lévy measure β which admits a density, say β_1, in the neighborhood of the origin, that is

$$\beta(dy) = \mathbf{1}_{|y|\leq 1}\beta_1(y)dy + \mathbf{1}_{|y|>1}\beta(dy).$$

We set $\beta_1(y) = 0$, for $|y| > 1$. Then

$$Z_t = \int_{(0,t]\times\mathbb{R}} y\ \tilde{N}_1(ds,dy) + \int_{(0,t]\times\mathbb{R}} y\ N_2(ds,dy), \tag{15}$$

where N_1, and N_2, are two independent Poisson point measures on $\mathbb{R}_+ \times \mathbb{R}$ with intensity measures equal, respectively, to $1_{|y|\leq 1}\beta_1(y)dy$, and $1_{|y|>1}\beta(dy)$, and $\tilde{N}_1$ is the compensated martingale measure of N_1.

We work here

- either under the assumptions of Subsection 1.1, i.e., $\mathbb{E}(|X_0|^2+|Z_T|^2)<+\infty$ and Lipschitz continuity of σ on $\mathbb{R}\times\mathcal{P}_2(\mathbb{R})$ endowed with the product of the Euclidean metric and the Vaserstein metric,
- or with the general assumptions of Subsection 1.2, i.e., no integrability conditions on Z and X_0, and Lipschitz continuity of σ on $\mathbb{R}\times\mathcal{P}(\mathbb{R})$ endowed with the product of the Euclidean metric and the modified Vaserstein metric d_1.

Propositions 2 and 6 ensure the existence of a probability measure solution P of (9). According to Proposition 7, there is a unique pathwise solution X (which is then unique in law), to the stochastic differential equation

$$X_t = X_0 + \int_{(0,t]\times\mathbb{R}} \sigma(X_{s-},P_s)\ y\tilde{N}_1(dy,ds) + \int_{(0,t]\times\mathbb{R}} \sigma(X_{s-},P_s)\ yN_2(dy,ds), \tag{16}$$

and, for $t\in[0,T]$, X_t is distributed according to the time marginal P_t.

Roughly speaking, our goal is to prove that, for each $t>0$, the probability measure P_t has a density with respect to the Lebesgue measure as long as the measure β, restricted to $[-1,1]$, has an infinite total mass due to an explosion of the density function $\beta_1(y)$ at 0. Indeed, we have in this case an accumulation of small jumps immediately after time 0, which will imply the absolute continuity of the law of X_t under suitable regularity assumptions on β_1.

For this purpose we develop a stochastic calculus of variations for diffusions with jumps driven by the Lévy process defined in (15). We thus generalize the approach developed by Bichteler and Jacod [4] (also, see Bismut [5]), who considered homogeneous processes with a jump measure equal to the Lebesgue measure. Here, the nonlinearity introduces an inhomogeneity in time, and the jump measure is much more general, which complicates the situation considerably and introduces additional difficulties. Graham and Méléard [11] developed similar techniques for a very specific stochastic differential equation related to the Kac equation. In that case, the jumps of the process were bounded. In our case, unbounded jumps are also allowed.

Our approach requires that we make the following standing assumptions on the coefficient $\sigma(x,\nu)$, and the Lévy density β_1:

Hypotheses (H):

1. The coefficient $\sigma(x,\nu)$ is twice differentiable in x.

2. There exist constants K_1, and K_2, such that, for each x, and ν,

$$|\sigma'_x(x,\nu)|\le K_1, \quad \text{and} \qquad |\sigma''_x(x,\nu)|\le K_2. \tag{17}$$

3. For each x, and ν,

$$\sigma(x,\nu)\neq 0. \tag{18}$$

Hypotheses ($\mathbf{H_1}$):

1. The function β_1 is twice continuously differentiable away from $\{0\}$.

2.

$$\int_{-1}^{1}\beta_1(y)dy=+\infty. \tag{19}$$

3. There exists a non-negative and non-constant function k of class C^1 on $[-1,1]$ such that $k(-1)=k(1)=0$, and such that

- $$\int_{-1}^{1} k^2(y)\beta_1(y)dy < +\infty, \tag{20}$$

- for all $y \in [-1,1]$,

$$\sup_{a\in[-K_1,K_1],\lambda\in[-1,1]} \frac{1}{|\lambda|}\left|\frac{\beta_1(y+\lambda(1+ay)k(y))}{\beta_1(y)}(1+\lambda(ak(y)+(1+ay)k'(y))-1\right| \le \frac{1}{2}, \tag{21}$$

- $$\sup_{a\in[-K_1,K_1]}\int_{-1}^{1}\left|\frac{\beta_1'(y)}{\beta_1(y)}(1+ay)k(y)+ak(y)+(1+ay)k'(y)\right|^2\beta_1(y)dy < +\infty, \tag{22}$$

- $$\sup_{a\in[-K_1,K_1]}\int_{-1}^{1}\sup_{\lambda\in[-1,1]}\left(\left|\frac{\beta_1''(y+\lambda(1+ay)k(y))}{\beta_1(y)}(1+\lambda(ak(y)+(1+ay)k'(y))\right|^2 k^2(y) + \left|\frac{\beta_1'(y+\lambda(1+ay)k(y))}{\beta_1(y)}\right|^2\right)k^2(y)\beta_1(y)dy < +\infty, \tag{23}$$

- for all $y \in [-1,1]$,

$$|k(y)| < \frac{1}{4(1+K_1)} \quad ; \quad |k'(y)| < \frac{1}{4(1+K_1)}. \tag{24}$$

The assumption ($\mathbf{H_1 3}$) on β_1 is obviously technical, but the assumption ($\mathbf{H_1 2}$) is essential, and cannot be avoided if one hopes to prove the absolute continuity result. The main example satisfying assumptions ($\mathbf{H_1}$) is the symmetric stable process with index $\alpha \in (0,2)$, for which $\beta(dy) = K\,dy/|y|^{1+\alpha}$, as developed in Section 3.

Theorem 9 *Consider the real-valued process X satisfying the nonlinear stochastic differential equation* (16)*. Assume that σ satisfies Hypotheses* ($\mathbf{H}$)*, and that β_1 satisfies Hypotheses* ($\mathbf{H_1}$)*. Then the law of the real-valued random variable X_T has a density with respect to the Lebesgue measure.*

The remainder of this section is devoted to the proof of Theorem 9 which will proceed through a series of lemmas and propositions. Our aim is to show that P_T has a density with respect to the Lebesgue's measure. Because of the compensated martingale term $\tilde{N}_1$ it would be natural to work with square integrable processes. But the finiteness of $\mathbb{E}\left(\sup_{t\le T}|X_t|^2\right)$ is not guaranteed because of the big jumps of N_2 . So, to develop the relevant stochastic calculus of variations in L^2, we use a trick defining $\mathbb{P}_T$ as the conditional law of (X_0, N_1, N_2) given (X_0, N_2^T), where N_2^T denotes the restriction of the measure N_2 to $[0,T]\times\mathbb{R}$. Thus, in what follows, the random variables considered being functions of (X_0, N_1, N_2) we may define their $\mathbb{E}_T$-expectations as the integral of the corresponding functions under $\mathbb{P}_T$. From now on, for notational simplicity, every statement concerning $\mathbb{E}_T$, or $\mathbb{P}_T$, holds almost everywhere under the law of (X_0, N_2^T), even if this fact is not mentioned explicitly. This conditioning allows us to use the same techniques as if the process X were square integrable. More precisely, given N_2^T, there are finitely many jump times and jump amplitudes of N_2 on $(0,T]$ and we will denote them by $(T_1, Y_1), \cdots, (T_k, Y_k)$.

Lemma 10

$$\mathbb{E}_T(\sup_{t\leq T} |X_t|^2) < +\infty. \tag{25}$$

Proof : As usual, we localize via $\tau_n = \inf\{t > 0, |X_t| \geq n\}$. Then, for $t \leq T$,

$$\begin{aligned}
\mathbb{E}_T\left(\sup_{s\leq t} |X_{t\wedge\tau_n}|^2\right) &\leq C\Bigg(|X_0|^2 + \int_0^{t\wedge\tau_n}\int_{-1}^1 |y|^2\Big(\mathbb{E}_T(|X_s|^2) + \sigma^2(0,P_s)\Big)\beta_1(y)dyds \\
&\quad + \sum_{i=1}^k |Y_i|^2\left(\mathbb{E}_T(|X_{T_i-}|^2) + \sigma^2(0,P_{T_i})\right)\mathbf{1}_{T_i\leq t\wedge\tau_n}\Bigg) \\
&\leq C\Bigg(|X_0|^2 + \sup_{u\in[0,T]}\sigma^2(0,P_u)\Big(1 + \sum_{i=1}^k |Y_i|^2\Big) + \int_0^{t\wedge\tau_n}\mathbb{E}_T(|X_s|^2)ds \\
&\quad + \sup_{i=1}^k |Y_i|^2\int_0^{t\wedge\tau_n}\mathbb{E}_T(|X_{s-}|^2)\int_{|y|>1}N_2(dy,ds)\Bigg).
\end{aligned}$$

At this point we apply Gronwall's Lemma in its generalized form (with respect to the measure $ds + \int_{|y|>1} N_2(dy, ds)$, see, for example, Ethier and Kurz [10], p. 498). The result then follows. ∎

Let us now explain our strategy to prove Theorem 9 before giving the technical details. We are going to prove that there exists an a.s. positive random variable DX_T such that $\mathbb{E}_T(DX_T) < +\infty$, and $\exists C\ \forall\phi \in C_c^\infty(\mathbb{R})$,

$$|\mathbb{E}_T(\phi'(X_T)DX_T)| \leq C\|\phi\|_\infty, \tag{26}$$

where $C_c^\infty(\mathbb{R})$ denotes the space of infinitely differentiable functions with compact support on the real line. Indeed, in the special case $DX_T = 1$ this inequality implies that the conditional law of X_T given (X_0, N_2^T) admits a density (see, for example, Nualart [17], p.79). If $DX_T \neq 1$, the law of X_T under $\mathbb{Q}_T = \frac{DX_T.\mathbb{P}_T}{\mathbb{E}_T(DX_T)}$ admits a density. But since $DX_T > 0$, a.s., then $\mathbb{Q}_T$ is equivalent to $\mathbb{P}_T$, and the conditional law of X_T given (X_0, N_2^T) still admits a density. Of course, this implies that the law P_T of X_T admits a density.

We will prove inequality (26) employing the stochastic calculus of variations. Consider perturbed paths of the process on the time interval $[0,T]$ and introduce a parameter $\lambda \in [-1,1]$, sufficiently close to 0. The perturbed Poisson measure N_1^λ will satisfy $N_1^0 = N_1$, and be such that, for a well chosen $\mathbb{P}_T$-martingale $(G_t^\lambda)_{t\leq T}$, the law of its restriction to $[0,T]$ under $G_T^\lambda.\mathbb{P}_T$ will be equal to the law of N_1 under $\mathbb{P}_T$. The process X^λ will be defined like X, only replacing N_1 by N_1^λ in the stochastic differential equation. Then, for sufficiently smooth functions ϕ, we will have

$$\mathbb{E}_T(\phi(X_T)) = \mathbb{E}_T(G_T^\lambda\phi(X_T^\lambda)). \tag{27}$$

Differentiating in λ at $\lambda = 0$, in a sense that we yet have to define, we will obtain

$$\mathbb{E}_T(\phi'(X_T)DX_T) = -\mathbb{E}_T(DG_T\phi(X_T)), \tag{28}$$

where $DX_T = \frac{d}{d\lambda}X_T^\lambda|_{\lambda=0}$, and $DG_T = \frac{d}{d\lambda}G_T^\lambda|_{\lambda=0}$. Then one easily deduces (26) with $C = \mathbb{E}_T(|DG_T|)$.

Let us describe the perturbation we are interested in. Let g be an increasing function of class $C_b^\infty(\mathbb{R})$, equal to x on $[-\frac{1}{2};\frac{1}{2}]$, equal to 1, for $x \geq 1$, and to -1, for $x \leq -1$. Note that $\|g\|_\infty \leq 1$, and that $g(x)x > 0$, for $x \in \mathbb{R}^*$.

Now, we define the predictable function $v : \Omega \times [0,T] \times [-1,1] \mapsto \mathbb{R}$ via the formula

$$v(s,y) = \mathbf{1}_{\{s>S\}}(1 + y\sigma_x'(X_{s-}, P_s))\ g(\sigma(X_{s-}, P_s))\ k(y) \tag{29}$$

where S is a stopping time that we are going to choose later on in order to ensure that $DX_T > 0$, a.s. (see the discussion before Proposition 15). It is easy to verify that the function $y \mapsto v(t,y)$ is of class C^1 on $[-1,1]$, and in what follows we will denote its derivative by $v'(t,y)$, Also, for every ω, t, and y,

$$|v(t,y)| \leq (1+K_1)k(y) \quad \text{and} \quad |v'(t,y)| \leq K_1 k(y) + (1+K_1)|k'(y)|. \tag{30}$$

For $\lambda \in [-1,1]$, let us introduce the perturbation function

$$\gamma^\lambda(t,y) = y + \lambda v(t,y). \tag{31}$$

We can easily check that, for every ω, and t, the map $y \mapsto \gamma^\lambda(t,y)$ is an increasing bijection from $[-1,1]$ into itself, since by (24) and (30), $|v'| \leq \frac{1}{2}$, and $k(-1) = k(1) = 0$.

Let us denote by N_1^λ the image measure of the Poisson point measure N_1 via the mapping γ^λ defined, for any Borel subset A of $[0,T] \times [-1,1]$, by the integral

$$N_1^\lambda(A) = \int \mathbf{1}_A(t, \gamma^\lambda(t,y)) N_1(dy,dt).$$

We also introduce the function

$$V^\lambda(s,y) = \frac{\beta_1(y + \lambda v(s,y))}{\beta_1(y)}(1 + \lambda v'(s,y)), \tag{32}$$

which appears below in the definition of the process G^λ (in Proposition 12). As a preliminary step, we obtain the following estimates concerning V^λ.

Lemma 11 *There exists a constant C such that, for almost all ω, and for all $s \in [0,T]$,*

$$\sup_{\lambda, y \in [-1,1]} \frac{1}{\lambda}|V^\lambda(s,y) - 1| \leq \frac{1}{2}\ , \tag{33}$$

$$\sup_{\lambda \in [-1,1]} \frac{1}{\lambda^2}\int_{-1}^1 |V^\lambda(s,y) - 1|^2 \beta_1(y)dy \leq C, \tag{34}$$

$$\sup_{\lambda \in [-1,1]} \frac{1}{\lambda^4}\int_{-1}^1 \left|V^\lambda(s,y) - 1 - \lambda\frac{d}{d\lambda}V^\lambda(s,y)/_{\lambda=0}\right|^2 \beta_1(y)dy \leq C. \tag{35}$$

Proof : Inequality (33) follows from (21). Also, one has

$$\frac{d}{d\lambda}V^\lambda(s,y) \;=\; v'(s,y)\frac{\beta_1(y+\lambda v(s,y))}{\beta_1(y)} + \frac{\beta_1'(y+\lambda v(s,y))}{\beta_1(y)}v(s,y)(1+\lambda v'(s,y)), \tag{36}$$

and

$$\frac{d^2}{d\lambda^2}V^\lambda(s,y) \quad = \quad v^2(s,y)(1+\lambda v'(s,y))\frac{\beta_1''(y+\lambda v(s,y))}{\beta_1(y)} + 2\frac{\beta_1'(y+\lambda v(s,y))}{\beta_1(y)}v(s,y)v'(s,y). \tag{37}$$

Since, for $\lambda \in [-1,1]$, we have estimates,

$$\left|\frac{V^\lambda(s,y)-1}{\lambda}\right|^2 \leq 2\left|\frac{d}{d\lambda}V^\lambda(s,y)/_{\lambda=0}\right|^2 + \frac{2}{\lambda^2}\left|V^\lambda(s,y)-1-\lambda\frac{d}{d\lambda}V^\lambda(s,y)/_{\lambda=0}\right|^2,$$

and

$$\left|V^\lambda(s,y)-1-\lambda\frac{d}{d\lambda}V^\lambda(s,y)/_{\lambda=0}\right|^2 \leq \frac{\lambda^4}{4}\sup_{\mu\in[-1,1]}\left|\frac{d^2}{d\mu^2}V^\mu(s,y)\right|^2,$$

one deduces (35) (resp. (34)) from (23), and (24)(resp. (22), (23), and (24)). ∎

We are now ready to introduce the promised earlier definition of the process G^λ.

Proposition 12 *(i) For every $\lambda \in [-1,1]$, the stochastic differential equation*

$$G_t^\lambda = 1 + \int_{(0,t]\times\mathbb{R}} G_{s-}^\lambda(V^\lambda(s,y)-1)\tilde{N}_1(dy,ds), \tag{38}$$

has a unique solution G^λ which is a strictly positive martingale under $\mathbb{P}_T$ and such that

$$\sup_{\lambda\in[-1,1]} \mathbb{E}_T\left(\sup_{t\leq T}|G_t^\lambda|^2\right) < +\infty. \tag{39}$$

(ii) The law of N_1^λ under $\mathbb{P}_T^\lambda = G_T^\lambda.\mathbb{P}_T$ is the same as the law of N_1 under $\mathbb{P}_T$.

Proof : *(i)* Thanks to (34), the stochastic integral $M_t^\lambda = \int_{(0,t]\times\mathbb{R}}(V^\lambda(s,y)-1)\tilde{N}_1(dy,ds)$ is well defined and is a $\mathbb{P}_T$ square integrable martingale. The unique solution to (38) is the exponential martingale $G_t^\lambda = \mathcal{E}(M^\lambda)_t = e^{M_t^\lambda}\prod_{0<s\leq t}(1+\Delta M_s^\lambda)e^{-\Delta M_s^\lambda}$ given by the Doléans-Dade formula.

Using (33), we remark that the jumps of M^λ are more than $-1/2$ so that G_t^λ is positive for each t. Moreover, using (34), Doob's inequality and Gronwall's Lemma, we deduce from (38) that (39) holds.

(ii) Let us denote $\mu(dy,dt) = \mathbf{1}_{|y|\leq 1}\beta_1(y)dydt$, and compute the image measure $\gamma^\lambda(V^\lambda.\mu)$. For a Borel subset A of $[0,T]\times[-1,1]$, we have

$$\begin{aligned}
\gamma^\lambda(V^\lambda.\mu)(A) &= \int \mathbf{1}_A(t,y+\lambda v(t,y))V^\lambda(t,y)\beta_1(y)dydt \\
&= \int \mathbf{1}_A(t,y')\frac{\beta_1(y')}{\beta_1(y)}(1+\lambda v'(t,y))\beta_1(y)\frac{dy'}{1+\lambda v'(t,y)}dt \\
&= \int \mathbf{1}_A(t,y')\beta_1(y')dy'dt = \mu(A),
\end{aligned} \tag{40}$$

where $y' = y + \lambda v(t,y)$. Hence

$$\gamma^\lambda(V^\lambda.\mu) = \mu. \tag{41}$$

Since N_1 is independent from (X_0, N_2^T), the compensator of N_1 under $\mathbb{P}_T$ is μ. By the Girsanov's theorem for random measures (cf. Jacod and Shiryaev [12], p. 157), its compensator under $\mathbb{P}_T^\lambda = G_T^\lambda.\mathbb{P}_T$ is $V^\lambda.\mu$ and thus, the compensator of $N_1^\lambda = \gamma^\lambda(N_1)$ is equal to $\gamma^\lambda(V^\lambda.\mu) = \mu$. We have thus proven that the compensator of N_1^λ under $\mathbb{P}_T^\lambda$ is μ, and the second assertion in the proposition follows. ∎

Next, we study the differentiability of G^λ with respect to the parameter λ, at $\lambda = 0$.

Proposition 13 *(i) The process*

$$DG_t = \int_{(0,t]\times\mathbb{R}} \frac{d}{d\lambda} V^\lambda(s,y)_{/\lambda=0} \tilde{N}_1(dy,ds) = \int_{(0,t]\times\mathbb{R}} \left(v'(s,y) + \frac{\beta_1'(y)}{\beta_1(y)} v(s,y) \right) \tilde{N}_1(dy,ds) \tag{42}$$

is well defined, and such that

$$\mathbb{E}_T(\sup_{t\leq T} |DG_t|^2) < +\infty. \tag{43}$$

(ii) The process G^λ is L^2-differentiable at $\lambda = 0$, with the derivative DG which is understood in the following sense:

$$\mathbb{E}_T\left(\sup_{t\leq T} |G_t^\lambda - 1 - \lambda DG_t|^2\right) = o(\lambda^2), \quad \text{a.s.}, \tag{44}$$

as λ tends to 0.

Proof : *(i)* Thanks to (36) and (22), for almost all ω, and all $s \in [0,T]$,

$$\int_{-1}^1 \left| \frac{d}{d\lambda} V^\lambda(s,y)_{/\lambda=0} \right|^2 \beta_1(y) dy \leq C. \tag{45}$$

Therefore the process DG_t is well defined and satisfies (43).

(ii) Moreover, one has

$$\begin{aligned}
&\mathbb{E}_T\left(\sup_{t\leq T} |G_t^\lambda - 1 - \lambda DG_t|^2\right) \\
\leq\ & C \int_{(0,t]\times\mathbb{R}} \mathbb{E}_T\left(\left(G_s^\lambda \left(V^\lambda(s,y) - 1\right) - \lambda \frac{d}{d\lambda} V^\lambda(s,y)_{/\lambda=0}\right)^2\right) \beta_1(y) dyds \\
\leq\ & C \int_{(0,t]\times\mathbb{R}} \mathbb{E}_T\left(\left(G_s^\lambda \left(V^\lambda(s,y) - 1 - \lambda \frac{d}{d\lambda} V^\lambda(s,y)_{/\lambda=0}\right)\right)^2\right) \beta_1(y) dyds \\
&+ \lambda^2 C \int_{(0,t]\times\mathbb{R}} \mathbb{E}_T\left(\left(\frac{d}{d\lambda} V^\lambda(s,y)_{/\lambda=0}\right)^2 (G_s^\lambda - 1)^2\right) \beta_1(y) dyds.
\end{aligned}$$

Now, according to (35) and (39), we obtain that

$$\int_{(0,t]\times\mathbb{R}} \mathbb{E}_T\left(\left(G_s^\lambda (V^\lambda(s,y) - 1 - \lambda \frac{d}{d\lambda} V^\lambda(s,y)_{/\lambda=0})\right)^2\right) \beta_1(y) dyds \leq C\lambda^4 t.$$

Furthermore, by (45),

$$\int_{(0,t]\times\mathbb{R}} \mathbb{E}_T\left(\left(\frac{d}{d\lambda}V^\lambda(s,y)_{/\lambda=0})\right)^2 (G_s^\lambda-1)^2\right)\beta_1(y)dyds \le \int_0^t \mathbb{E}_T((G_s^\lambda-1)^2)ds,$$

and using (34) and (39), we may show that, for each $t\le T$,

$$\mathbb{E}_T((G_t^\lambda-1)^2) = \int_{(0,t]\times\mathbb{R}} \mathbb{E}_T((G_s^\lambda(V^\lambda(s,y)-1))^2)\beta_1(y)dyds \le C\lambda^2.$$

This concludes the proof. ∎

In the next step we define the perturbed stochastic differential equation. Let us recall that the probability measures P_t are fixed and are considered as time-dependent parameters. Thus the process X is a function $F_P(X_0,N_1,N_2)$ of the triplet (X_0,N_1,N_2).

Define $X^\lambda := F_P(X_0,N_1^\lambda,N_2)$. Hence, using Proposition 12 *(ii)*, the law of X^λ under $\mathbb{P}_T^\lambda$ is equal to the law of X under $\mathbb{P}_T$. A simple computation shows that X^λ is a solution of the stochastic differential equation

$$\begin{aligned}
X_t^\lambda &= X_0 + \int_{(0,t]\times\mathbb{R}} y\,\sigma(X_{s-}^\lambda,P_s)\,\Big(N_1^\lambda(dy,ds)-\beta_1(dy)ds\Big) + \int_{(0,t]\times\mathbb{R}} y\,\sigma(X_{s-}^\lambda,P_s)N_2(dy,ds)\\
&= X_0 + \int_{(0,t]\times\mathbb{R}} (y+\lambda v(s,y))\,\sigma(X_{s-}^\lambda,P_s)\,\Big(N_1(dy,ds)-V^\lambda(s,y)\beta_1(dy)ds\Big)\\
&\qquad + \int_{(0,t]\times\mathbb{R}} y\,\sigma(X_{s-}^\lambda,P_s)N_2(dy,ds), \qquad\qquad (\text{ since } \gamma^\lambda(V^\lambda.\mu)=\mu)\\
&= X_0 + \int_{(0,t]\times\mathbb{R}} y\sigma(X_{s-}^\lambda,P_s)\tilde N_1(dy,ds) + \lambda\int_{(0,t]\times\mathbb{R}} \sigma(X_{s-}^\lambda,P_s)v(s,y)\tilde N_1(dy,ds)\\
&+ \int_{(0,t]\times\mathbb{R}} y\,\sigma(X_{s-}^\lambda,P_s)N_2(dy,ds) - \int_{(0,t]\times\mathbb{R}} (y+\lambda v(s,y))\sigma(X_{s-}^\lambda,P_s)(V^\lambda(s,y)-1)\beta_1(y)dyds.
\end{aligned} \tag{46}$$

Using (20) for the second term, the fact that $\int_{-1}^1(y^2+k^2(y))\beta_1(y)dy<+\infty$, (34), and the Cauchy-Schwarz inequality for the last term, we easily prove that equation (46) has a unique pathwise solution.

Let us now show that X^λ is differentiable in λ, at $\lambda=0$, in the L^2-sense. More precisely we have the following

Proposition 14

$$(i) \qquad \mathbb{E}_T\left(\sup_{t\le T}|X_t^\lambda - X_t|^4\right) \le C\lambda^4. \tag{47}$$

$$(ii) \qquad \mathbb{E}_T\left(\sup_{t\le T}|X_t^\lambda - X_t-\lambda DX_t|^2\right) = o(\lambda^2), \tag{48}$$

as λ tends to 0, where DX is a solution of the affine stochastic differential equation

$$DX_t = \int_{(0,t]\times\mathbb{R}} y\sigma'_x(X_{s-},P_s)DX_{s-}\;\tilde N_1(dy,ds) + \int_{(0,t]\times\mathbb{R}} \sigma(X_{s-},P_s)v(s,y)\tilde N_1(dy,ds)$$

$$+\int_{(0,t]\times\mathbb{R}} y\sigma'_x(X_{s-},P_s)DX_{s-}\ N_2(dy,ds)-\int_{(0,t]\times[-1,1]} y\ \sigma(X_{s-},P_s)(\beta_1(y)v'(s,y)+\beta'_1(y)v(s,y))dyds. \tag{49}$$

Proof : In order to prove assertion *(i)*, we need the following moment estimate :

$$\sup_{\lambda\in[-1,1]}\mathbb{E}_T\left(\sup_{t\leq T}|X_t^\lambda|^4\right)<+\infty. \tag{50}$$

It relies on the following classical estimation (see, [4], Lemme (A.14)):

$$\mathbb{E}_T\left(\left(\int_{(0,t]\times\mathbb{R}} H_s\rho(y)\tilde{N}_1(dy,ds)\right)^4\right)\leq C\left(\left(\int_{-1}^1\rho^2(y)\beta_1(y)dy\right)^2+\int_{-1}^1\rho^4(y)\beta_1(y)dy\right)$$
$$\times\int_0^t\mathbb{E}_T\left(\sup_{u\leq s}|H_u|^4\right)ds, \tag{51}$$

for any predictable process H, and any measurable function $\rho:[-1,1]\mapsto\mathbb{R}$ such that the right-hand side is finite. Conditioning by N_2^T, the times and the amplitudes of jumps of N_2 on $(0,t]$ are given by $(T_1,Y_1),\cdots,(T_k,Y_k)$, and

$$\mathbb{E}_T\left(\left|\int_0^t y\sigma(X_{s-}^\lambda,P_s)N_2(dy,ds)\right|^4\right)\leq C\sum_{i=1}^k Y_i^4\left(\mathbb{E}_T(|X_{T_i-}^\lambda|^4)+\sigma^4(0,P_{T_i})\right)$$
$$\leq C\sup_{i=1}^k|Y_i|^4\left(\sup_{u\in[0,T]}\sigma^4(0,P_u)+\int_0^t\int_{|y|>1}\mathbb{E}_T(|X_{s-}^\lambda|^4)N_2(dy,ds)\right).$$

Applying (51) with $\rho(y)=y$, and $\rho(y)=k(y)$, (34) and Gronwall's Lemma with respect to the measure $ds+\int_{|y|>1}N_2(dy,ds)$, we easily check (50) and deduce

$$\sup_{\lambda\in[-1,1]}\mathbb{E}_T\left(\sup_{t\leq T}|\sigma(X_t^\lambda,P_t)|^4\right)<+\infty.$$

Now, we write $X_t^\lambda-X_t$ using (16) and (46). Assertion *(i)* is obtained following an analogous argument.

To prove *(ii)* we need to isolate the term $Z_t^\lambda=X_t^\lambda-X_t-\lambda DX_t$, and as in [4], Theorem (A.10), we write

$$\begin{aligned}
&X_t^\lambda-X_t-\lambda DX_t\\
=\ &\int_{(0,t]\times\mathbb{R}} y\ Z_{s-}^\lambda\sigma'_x(X_{s-},P_s)(\tilde{N}_1(dy,ds)+N_2(dy,ds))\\
&+\int_{(0,t]\times\mathbb{R}} y\left(\sigma(X_{s-}^\lambda,P_s)-\sigma(X_{s-},P_s)-\sigma'_x(X_{s-},P_s)(X_{s-}^\lambda-X_{s-})\right)(\tilde{N}_1(dy,ds)+N_2(dy,ds))\\
&+\int_{(0,t]\times\mathbb{R}}\lambda v(s,y)\left(\sigma(X_{s-}^\lambda,P_s)-\sigma(X_{s-},P_s)\right)\tilde{N}_1(dy,ds)\\
&-\int_{(0,t]\times\mathbb{R}} y\ \sigma(X_s,P_s)\left(V^\lambda(s,y)-1-\lambda\frac{d}{d\lambda}V^\lambda(s,y)_{/\lambda=0}\right)\beta_1(y)dyds\\
&-\int_{(0,t]\times\mathbb{R}} y\ \left(\sigma(X_s^\lambda,P_s)-\sigma(X_s,P_s)\right)(V^\lambda(s,y)-1)\beta_1(y)dyds\\
&-\int_{(0,t]\times\mathbb{R}}\lambda\ v(s,y)\sigma(X_s^\lambda,P_s)(V^\lambda(s,y)-1)\beta_1(y)dyds.
\end{aligned}$$

Under Hypotheses (**H**) and (**H**$_1$) all the integral terms, except the first one, are of order λ^2. Indeed, for the second term, we use Taylor's expansion of σ, and (47); for the third term, we use (20), and (47); for the fourth we use the Cauchy-Schwarz inequality, and (35); for the fifth term, we use the Cauchy-Schwarz inequality, (34), and (47); for the sixth term, we use Cauchy-Schwarz inequality, (34), and (20). Then, as previously, using Gronwall's Lemma for the conditional expectation, we obtain the result. ∎

The term DX_t requires our special attention. Observe that, after integration by parts (in the variable y), the last term in (14) cancels the compensated part of

$$\int_{(0,t]\times\mathbb{R}} \sigma(X_{s-}, P_s) v(s,y) \tilde{N}_1(dy, ds),$$

and one obtains

$$DX_t = \int_0^t DX_{s-} dK_s + L_t \tag{52}$$

where

$$K_t = \int_{(0,t]\times\mathbb{R}} y\ \sigma'_x(X_{s-}, P_s)\tilde{N}_1(dy, ds) + \int_{(0,t]\times\mathbb{R}} y\ \sigma'_x(X_{s-}, P_s) N_2(dy, ds), \tag{53}$$

$$L_t = \int_{(0,t]\times\mathbb{R}} \sigma(X_{s-}, P_s)\ v(s,y)\ N_1(dy, ds). \tag{54}$$

As in [12], Theorem 4.61, p. 59, or in [4], we can solve (52) explicitely. The jumps ΔK_s are of the form $y\sigma'_x(X_s, P_s)$. Thus $1+\Delta K_s$ may be equal to 0 and then the Doleans-Dade exponential

$$\mathcal{E}(K)_t = e^{K_t}\Pi_{0<s\le t}(1+\Delta K_s)e^{-\Delta K_s}$$

vanishes from the first time when $\Delta K_s = -1$. We follow [4] to show that $DX_T \neq 0$, but the strict positivity (which has not been proved in the latter) necessitates a careful analysis.

Let us define the sequence of stopping times $S_1 = \inf\{t > 0, \Delta K_t \le -1\}$, $S_k = \inf\{t > S_{k-1}, \Delta K_t \le -1\}$, $S_0 = 0$. Since σ'_x is bounded, there is a finite number of big jumps on the time interval $[0,T]$, so that there exists an n such that $S_n \le T < S_{n+1} = \infty$, and $\mathbb{P}(S_n = T) = 0$.

Solving equation (52) gives

$$\begin{aligned} DX_t = \quad & \mathcal{E}(K - K^{S_k})_t \Big(DX_{S_k} + \int_{(S_k,t]} (1+\Delta K_s)^{-1} \mathcal{E}(K-K^{S_k})^{-1}_{s-} dL_s \Big) \\ & \text{if } S_k \le t < S_{k+1},\ \text{and } t \le T, \end{aligned} \tag{55}$$

where $K_t^{S_k} = K_{S_k\wedge t}$. In particular,

$$DX_T = \mathcal{E}(K-K^{S_n})_T \Big(DX_{S_n} + \int_{(S_n,T]} (1+\Delta K_s)^{-1} \mathcal{E}(K-K^{S_n})^{-1}_{s-} dL_s \Big).$$

Because of the definition of S_n, the exponential martingale $\mathcal{E}(K-K^{S_n})_s$ is non-negative on $[S_n, T]$. If the perturbation v did not vanish before time S_n, it would not be clear how to control the sign of DX_{S_n}. That is why we choose $S = S_n$ in (29) :

$$v(s,y) = \mathbf{1}_{\{s>S_n\}}(1 + y\ \sigma'_x(X_{s-}, P_s))\ k(y)\ g(\sigma(X_{s-}, P_s))$$

so that $DX_{S_n} = 0$. For this choice, we obtain

Proposition 15 *We have $DX_T > 0$, almost surely.*

Proof : One has $DX_T = \mathcal{E}(K - K^{S_n})_T Y_n$, where $Y_n = \int_{(S_n,T]}(1+\Delta K_s)^{-1}\mathcal{E}(K - K^{S_n})_{s-}^{-1} dL_s$.

Since N_1 and N_2 are independent, the sets of jumps are almost surely distinct and then $1 + \Delta K_s$ can be replaced by 1 every time the jump of K comes from a jump of N_2. So,

$$Y_n = \int_{(S_n,T]\times[-1,1]} h_n(s,y) N_1(ds,dy),$$

where

$$\begin{aligned} h_n(s,y) &= \mathcal{E}(K - K^{S_n})_{s-}^{-1}(1 + y\sigma'_x(X_{s-},P_s))^{-1} v(s,y)\sigma(X_{s-},P_s) \\ &= \mathcal{E}(K - K^{S_n})_{s-}^{-1} k(y)\sigma(X_{s-},P_s) g(\sigma(X_{s-},P_s)) \geq 0. \end{aligned}$$

Let us consider the set $A_n = \{(\omega,s,y), h_n(\omega,s,y) > 0\}$ and define the stopping time $\tau = \inf\{t > S_n, \int_{(S_n,t]\times[-1,1]} h_n(s,y)N_1(ds,dy) > 0\} = \inf\{t > S_n, \int_{(S_n,t]\times[-1,1]} \mathbf{1}_{A_n}(s,y)N_1(ds,dy) > 0\}$.

Using the definitions of v and S_{n+1}, one knows that, if $S_n(\omega) < s \leq T \wedge S_{n+1}(\omega)$, then

$$(\omega,s,y) \in A_n \Leftrightarrow \sigma(X_{s-}(\omega),P_s) \neq 0,$$

which is always the case in view of Hypothesis (**H**).

On the other hand, $\int_{(S_n,\tau]\times[-1,1]} \mathbf{1}_{A_n}(s,y)N_1(ds,dy) \leq 1$, so $\mathbb{E}\left(\int_{(S_n,\tau]\times[-1,1]} \mathbf{1}_{A_n}(s,y)N_1(ds,dy)\right) \leq 1$ and, for ω in a set of probability 1,

$$\int_{(S_n,\tau]\times[-1,1]} \mathbf{1}_{A_n}(s,y)\beta_1(y)dyds < +\infty.$$

These two remarks, and the fact the $\int_{-1}^{1}\beta_1(y)dy = +\infty$, imply that $\tau = S_n$, almost surely. So Y_n is strictly positive. Therefore DX_T is strictly positive as well and the proof is complete. ∎

We are now in a position to complete the proof of Theorem 9.

Proof of Theorem 10 : For $\phi \in C_b^\infty(\mathbb{R})$, we differentiate the expression (27) in the L^2-sense with respect to λ, at $\lambda = 0$, and hence, we obtain the "integration-by-parts" formula (28). Then, since $\mathbb{E}_T|DG_T| < +\infty$, we obtain (26), which concludes the proof. ∎

3 The case of a symmetric stable driving process Z

In this section, we assume that the real-valued driving process Z is a symmetric stable process with index $\alpha \in (0,2)$, i.e., Z is given by (15) with $\beta(dy) = \frac{K}{|y|^{1+\alpha}}dy$, where $K > 0$ is a normalization constant. The generator of this process is the fractional Laplacian (or, fractional symmetric derivative) of order α on $\mathbb{R}$:

$$D_x^\alpha f(x) = K\int_{\mathbb{R}} \left(f(x+y) - f(x) - \mathbf{1}_{\{|y|\leq 1\}} f'(x)y\right)\frac{dy}{|y|^{1+\alpha}}.$$

This operator may be defined alternatively via the Fourier transform $\mathcal{F}$:

$$D_x^\alpha v(x) = K'\mathcal{F}^{-1}\Big(|\xi|^\alpha \mathcal{F}(v)(\xi)\Big)(x), \text{ with } K' > 0.$$

When $\sigma : \mathbb{R}\times\mathcal{P}(\mathbb{R}) \to \mathbb{R}$ satisfies Hypotheses (**H**), it is possible to explicitly calculate the adjoint $\mathcal{L}^*[\nu]$ involved in the nonlinear Fokker-Planck equation (13). For smooth functions $\varphi, \psi : \mathbb{R} \mapsto \mathbb{R}$, one has

$$\int_{\mathbb{R}} \mathcal{L}[\nu]\varphi(x)\psi(x)dx = K \int_{\mathbb{R}^2} \left(\varphi(x + s(x)y) - \varphi(x) - \mathbf{1}_{\{|y|\leq 1\}} s(x)y\varphi'(x)\right) \frac{dy}{|y|^{1+\alpha}}\psi(x)dx,$$

where $s(x) = \sigma(x, \nu)$. Setting $z = -s(x)y$, and observing that $\int_{\mathbb{R}} \left(\mathbf{1}_{\{|z|\leq s(x)\}} - \mathbf{1}_{\{|z|\leq 1\}}\right) \frac{zdz}{|z|^{1+\alpha}} = 0$, one gets

$$\int_{\mathbb{R}} \left(\varphi(x + s(x)y) - \varphi(x) - \mathbf{1}_{\{|y|\leq 1\}} s(x)y\varphi'(x)\right) \frac{dy}{|y|^{1+\alpha}}$$
$$= \int_{\mathbb{R}} \left(\varphi(x - z) - \varphi(x) + \mathbf{1}_{\{|z|\leq 1\}} z\varphi'(x)\right) |s(x)|^{\alpha} \frac{dz}{|z|^{1+\alpha}}.$$

Since

$$\int_{\mathbb{R}} \varphi(x - z)[|s|^{\alpha}\psi](x)dx = \int_{\mathbb{R}} \varphi(x)[|s|^{\alpha}\psi](x + z)dx,$$

and

$$\int_{\mathbb{R}} \varphi'(x)[|s|^{\alpha}\psi](x)dx = -\int_{\mathbb{R}} \varphi(x)[|s|^{\alpha}\psi]'(x)dx,$$

invoking Fubini's theorem one concludes that

$$\int_{\mathbb{R}} \mathcal{L}[\nu]\varphi(x)\psi(x)dx = K \int_{\mathbb{R}} \varphi(x) \int_{\mathbb{R}} \left([|s|^{\alpha}\psi](x + z) - [|s|^{\alpha}\psi](x) - \mathbf{1}_{\{|z|\leq 1\}} z[|s|^{\alpha}\psi]'(x)\right) \frac{dz}{|z|^{1+\alpha}}dx.$$

Therefore

$$\mathcal{L}^*[\nu]\psi(x) = D^{\alpha}_x(|\sigma(., \nu)|^{\alpha}\psi(.))(x).$$

Moreover, the absolute continuity result given in Theorem 9 permits us to prove existence of a function solution to the nonlinear Fokker-Planck equation.

Theorem 16 *Let $m \in \mathcal{P}(\mathbb{R})$, and $\alpha \in (0, 2)$. Assume that the function $\sigma(x, \nu)$ satisfies hypotheses* (**H**) *and is Lipschitz continuous in its second variable when $\mathcal{P}(\mathbb{R})$ is endowed with the modified Vaserstein metric d_1. Then, there exists a function $(t, x) \in (0, T] \times \mathbb{R} \mapsto p_t(x) \in \mathbb{R}_+$ such that, for each $t \in (0, T]$, $\int_{\mathbb{R}} p_t(x)dx = 1$ and, in the weak sense,*

$$\begin{cases} \partial_t p_t(x) = D^{\alpha}_x(|\sigma(., p_t)|^{\alpha}p_t(.))(x) \\ \lim_{t\to 0^+} p_t(x)dx = m(dx) \text{ for the weak convergence,} \end{cases} \tag{56}$$

where, by a slight abuse of notation, $\sigma(., p_t)$ stands for $\sigma(., p_t(y)dy)$.

Proof : Existence of a measure solution $(P_t)_{t\in[0,T]}$ to the nonlinear Fokker-Planck equation follows from Propositions 6 and 7. So to conclude the proof, it is enough to exhibit a perturbation function $k(y)$ satisfying hypotheses ($\mathbf{H}_1$) with $\beta_1(y) = \mathbf{1}_{\{|y|\leq 1\}} \frac{K}{|y|^{1+\alpha}}$. Then, by Theorem 9, for each $t \in (0, T]$, we have $P_t = p_t(x)dx$.

For $\gamma > \frac{\alpha}{2}$, and $\varepsilon \in (0, 1/2)$, let k_{ε} denote the even function on $[-1, 1]$ defined by

$$k_{\varepsilon}(y) = \begin{cases} y^{1+\gamma}, & \text{for } y \in [0, \varepsilon], \\ \varepsilon^{1+\gamma} + (1 + \gamma)\varepsilon^{\gamma}(y - \varepsilon) - (1 + c)(y - \varepsilon)^{1+\gamma}, & \text{for } y \in [\varepsilon, 2\varepsilon] \\ (1 + \gamma - c)\varepsilon^{1+\gamma} - c(1 + \gamma)\varepsilon^{\gamma}(y - 2\varepsilon), & \text{for } y \in [2\varepsilon, 1], \end{cases},$$

where $c = \frac{(1+\gamma)\varepsilon}{(1+\gamma)-\varepsilon(1+2\gamma)}$, so that $k_\varepsilon(1) = 0$. The function k_ε is non-negative and C^1 on $[0,1]$, satisfies (20) and, $\forall y \in [-1,1]$,

$$k_\varepsilon(y) \leq (2+\gamma)\varepsilon^{1+\gamma}, \quad |k'_\varepsilon(y)| \leq (1+\gamma)\max(1,c)\varepsilon^\gamma, \quad \text{and} \quad \frac{k_\varepsilon(y)}{|y|} \leq (1+\gamma)\varepsilon^\gamma. \tag{57}$$

In particular, for small enough ε, k_ε satisfes (21). Since

$$\begin{aligned}\left|\frac{\beta'_1(y)}{\beta_1(y)}(1+ay)k_\varepsilon(y) + ak_\varepsilon(y) + (1+ay)k'_\varepsilon(y)\right|^2 \beta_1(y) &\leq C\left[\frac{k_\varepsilon^2(y)}{y^2} + k_\varepsilon^2(y) + (k'_\varepsilon)^2(y)\right]\beta_1(y) \\ &\sim C'|y|^{-(1+\alpha-2\gamma)},\end{aligned}$$

in the neighbourhood of 0, (22) is satisfied as well. In the same way, in the neighbourhood of 0,

$$\begin{aligned}\sup_{a\in[-K_1,K_1],\lambda\in[-1,1]} \Bigg(&\left|\frac{\beta''_1(y+\lambda(1+ay)k_\varepsilon(y))}{\beta_1(y)}(1+\lambda(ak_\varepsilon(y)+(1+ay)k'_\varepsilon(y))\right|^2 k_\varepsilon^2(y) \\ &+ \left|\frac{\beta'_1(y+\lambda(1+ay)k_\varepsilon(y))}{\beta_1(y)}\right|^2\Bigg) k_\varepsilon^2(y)\beta_1(y) \leq C\left(\frac{|y|^{2+2\gamma}}{y^4} + \frac{1}{y^2}\right)|y|^{1+2\gamma-\alpha},\end{aligned}$$

and (23) is satisfied.

Finally, for $a \in [-K_1, K_1]$ and $y, \lambda \in [-1,1]$, by (57), for $\varepsilon < ((1+K_1)(1+\gamma))^{-1/\gamma}$,

$$\begin{aligned}&\frac{1}{|\lambda|}\left|\frac{\beta_1(y+\lambda(1+ay)k_\varepsilon(y))}{\beta_1(y)}(1+\lambda(ak_\varepsilon(y)+(1+ay)k'_\varepsilon(y))-1\right| \\ =&\frac{1}{|\lambda|}\frac{|1-|1+\lambda(1+ay)\frac{k_\varepsilon(y)}{y}|^{1+\alpha}+\lambda(ak_\varepsilon(y)+(1+ay)k'_\varepsilon(y))|}{|1+\lambda(1+ay)\frac{k_\varepsilon(y)}{y}|^{1+\alpha}} \\ \leq&\frac{1}{(1-(1+K_1)(1+\gamma)\varepsilon^\gamma)^{1+\alpha}}\Big[(1+\alpha)(1+(1+K_1)(1+\gamma)\varepsilon^\gamma)^\alpha(1+K_1)(1+\gamma)\varepsilon^\gamma \\ &\qquad + K_1(2+\gamma)\varepsilon^{1+\gamma} + (1+K_1)(1+\gamma)\max(1,c)\varepsilon^\gamma\Big],\end{aligned}$$

and (24) is also satisfied for small enough ε. ∎

Remark 17 One of the motivations for our work was to generalize the probabilistic approximation of the porous medium equation

$$\partial_t p_t(x) = D_x^2(p_t^q(x)), \qquad q > 1, \tag{58}$$

developed, among others, by Jourdain [9] to the fractional case where D_x^2 is replaced by D_x^α. The equation (58), which describes percolation of gases through porous media, and which is usually derived by combining the power type equation of state relating pressure to gas density p, conservation of mass law, and so called Darcy's law describing the local gas velocity as the gradient of pressure, goes back, at least, to the 1930's (see, e.g., Muskat [16]). The major steps in the development of the mathematical theory of (58) were the discovery of the family of its self-similar solutions by Barenblatt (see, [1], and [2]) who obtained this equation in the context of heat propagation at the initial stages of a nuclear explosion, and an elegant uniqueness result for (58) proved by Brézis and Crandall [7]. A summary of some of the newer developments in the area of the standard porous medium equation can be found in a survey by Otto [18].

However, in a number of recent physical papers, an argument was made that some of the fractional scaling observed in flows-in-porous-media phenomena cannot be modeled in the framework of (58). In particular, Meerschaert, Benson and Baeumer [15] replace the Laplacian D_x^2 in (58) by the fractional Laplacian D_x^α while considering the linear case ($q = 1$) in a multidimensional case of anomalous (mostly geophysical) diffusion in porous media, while Park, Kleinfelter and Cushman [19] continue in this tradition and derive scaling laws and (linear) Fokker-Planck equations for 3-dimensional porous media with fractal mesoscale.

On the other hand, Tsallis and Bukman [23] suggest an alternative approach to the anomalous scaling problem (in porous media, surface growth, and certain biological phenomena) and consider an equation of the general form

$$\partial_t p_t^r(x) = -D_x(F(x)p_t^r(x)) + D_x^2(p_t^q(x)), \qquad r, q \in \mathbb{R}, \tag{59}$$

where $F(x)$ is an external force. The authors manage to find exact solutions for this class of equations using ingeniously the concept of Renyi (-Tsallis) entropy but, significantly, suggest in the conclusion of their paper that it would be desirable to develop physically significant models for which further unification can possibly be achieved by considering the generic case of a *nonlinear* Fokker-Planck-like equation with *fractional* derivatives. This is what we endeavored to do taking as our criterion of "physicality" the existence of an approximating interacting particle scheme. For the most obvious, simply-minded generalization, $\partial_t p_t(x) = D_x^\alpha(p_t^q(x))$, that physical interpretation seems to be missing, or, at least, we were unable to produce it and, as a result, our study lead us to settle on an equation like (56). Indeed for

$$\sigma(x,\nu) = (g_\varepsilon * \nu(x))^s \text{ with } \varepsilon > 0,\ g_\varepsilon(x) = \frac{1}{\sqrt{2\pi\varepsilon}} e^{-\frac{x^2}{2\varepsilon}} \text{ and } s > 0,$$

(56) writes $\partial_t p_t = D_x^\alpha((g_\varepsilon * p_t)^{\alpha s} p_t)$ which, for now, we are viewing as a "physically justifiable", fractional, *and* strongly nonlinear "extension" of the classical porous medium equation. Of course, this is only the beginning of the effort to understand these types of models.

References

[1] Barenblatt, G.I., On some unsteady motions of fluids and gases in a porous medium, Prikl. Mat. Mekh 16:67-78, 1952.

[2] Barenblatt, G.I., *Scaling, Self-similarity, and Intermediate Asymptotics*, Cambridge University Press, 1996.

[3] Bhatt, A.G. and Karandikar, R.L., Invariant measures and evolution equations for Markov processes characterized via martingale problems, Ann. Probab. 21(4):2246-2268, 1993.

[4] Bichteler, K. and Jacod, J., Calcul de Malliavin pour les diffusions avec sauts: Existence d'une densité dans le cas unidimensionnel. Séminaire de Probabilités XVII, Lect. Notes in Math. 986, Berlin-Heidelberg- New York: Springer, 132-157, 1983.

[5] Bismut, J.M., Calcul des variations stochastiques et processus de sauts, Z. Wahrsch. Verw. Gebiete 63(2):147-235, 1983.

[6] Bucklew, J.A. and Wise, G.L., Multidimensional asymptotic quantization theory with rth power distortion measures, IEEE Trans. Inform. Theory, 28(2):239-247, 1982.

[7] Brézis, H. and Crandall, M.G., Uniqueness of solutions of the initial-value problem for $u_t - \Delta\varphi(u) = 0$, J. math. pures et appl. 58:153-163, 1979.

[8] Del Barrio, E., Giné, E. and Utzet, F., Asymptotics for L_2 functionals of the empirical quantile process, with applications to tests of fit based on weighted Wasserstein distances, Bernoulli 11(1):131-189, 2005.

[9] Jourdain, B. , Probabilistic approximation for a porous medium equation, Stochastic Processes and their Applications 89:81-99, 2000.

[10] Ethier, S.N. and Kurtz, T.G., *Markov processes, Characterization and convergence*, Wiley 1986.

[11] Graham, C. and Méléard, S., Existence and Regularity of a solution of a Kac equation without cutoff using the stochastic calculus of variations, Commun. Math. Phys. 205(3):551-69, 1999.

[12] Jacod, J. and Shiryaev, A., *Limit theorems for stochastic processes*, Springer 1987.

[13] Jourdain, B., Méléard, S. and Woyczynski, W.A., Probabilistic approximation and inviscid limits for 1-D fractional conservation laws, Bernoulli, 11(4):689-714, 2005

[14] Jourdain, B., Méléard, S. and Woyczynski, W.A., A probabilistic approach for nonlinear equations involving fractional Laplacian and singular operator, Potential Analysis 23(1):55-81, 2005.

[15] Meerschaert, M.M., Benson, D.A. and Baeumer, B., Multidimensional advection and fractional dispersion, Phys. Rev. E, 59: 5026-5028, 1999.

[16] Muscat, M., *The Flow of Homogeneous Fluids Through Porous Media*, McGraw-Hill, New York 1937.

[17] Nualart, D., *The Malliavin calculus and related topics*, Springer 1995.

[18] Otto, F., The geometry of dissipative evolution equations: the porous medium equation, Commun. in Partial Differential Equations, 26: 101-174, 2001.

[19] Park, M., Kleinfelter, N. and Cushman, J.H., Scaling laws and Fokker-Planck equations for 3-dimensional porous media with fractal mesoscale, SIAM J. Multiscale Model. Simul. 4: 1233-1244, 2005.

[20] Protter, P., Stochastic integration and differential equations. Second edition. *Applications of Mathematics (New York), 21. Stochastic Modelling and Applied Probability. Springer-Verlag, Berlin*, 2004.

[21] Rachev, S.T. and Rüschendorf, L., Mass transportation problems. Vol. I and II, *Probability and its Applications (New York). Springer-Verlag, New York*, 1998.

[22] Sznitman, A.S., Topics in propagation of chaos. *Ecole d'été de probabilités de Saint-Flour XIX - 1989, Lect. Notes in Math. 1464, Springer-Verlag, 1991.*

[23] Tsallis, C. and Bukman, D.J., Anomalous diffusion in the presence of external forces: Exact time-dependent solutions and their thermostatistical basis, Phys. Rev. E 54(3):2197-2200, 1996.